PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/53, 9/02, 15/82, A01H 5/00

A2

(11) International Publication Number: **WO 98/50558**

(43) International Publication Date: 12 November 1998 (12.11.98)

(21) International Application Number: PCT/EP98/02830

(22) International Filing Date: 7 May 1998 (07.05.98)

(30) Priority Data:
1006000 7 May 1997 (07.05.97) NL

(71) Applicant *(for all designated States except US)*: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE [BE/BE]; Rijvisschestraat 118, bus 1, B–9052 Zwijnaarde–Poperinge (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BAUW, Guy, Jerome, Corneel [BE/BE]; Blokweg 23, B–8972 Proven–Poperinge (BE). DAVEY, Mark, William [GB/BE]; Patijntjestraat 68, B–9000 Gent (BE). OSTERGAARD, Jens [DK/BE]; Ekkergemstraat 165, B–9000 Gent (BE). VAN MONTAGU, Marc, Charles, Ernest [BE/BE]; De Stassartstraat 120, B–1050 Brussel (BE).

(74) Agent: VAN SOMEREN, Petronella, Francisca, Hendrika, Maria; Arnold & Siedsma, Sweelinckplein 1, NL–2517 GK The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PRODUCTION OF ASCORBIC ACID IN PLANTS

(57) Abstract

The present invention relates to a polynucleotide in isolated form, which polynucleotide codes for a protein with the activity of the enzyme L–galactono–γ–lactone dehydrogenase, which polynucleotide comprises at least the L–galactono–γ–lactone dehydrogenase activity–determining parts of the coding part of the nucleotide sequence, which is shown in Fig. 3, or a sequence derived therefrom on the basis of the degeneration of the genetic code. The invention further relates to the use of the polynucleotide in the production of transgenic plants, plant cells or other eukaryotic cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

PRODUCTION OF ASCORBIC ACID IN PLANTS

The present invention relates to a polynucleotide, in particular a cDNA, which codes for L-galactono-γ-lactone dehydrogenase (GLDase), an enzyme involved in the biosynthesis of ascorbic acid (vitamin C) in plants. The invention further relates to the use of this cDNA for the synthesis of the enzyme and for the production of transgenic plant and animal cells, plant tissues and plants producing the enzyme.

Ascorbic acid is synthesized in all higher plants and in almost all higher animals, with the exception of humans and other primates, the guinea pig and a number of birds. Opinions differ concerning the presence of ascorbic acid in micro-organisms. It appears to be present in a number of yeasts, although there are also reports which suggest that ascorbic acid analogues are found in micro-organisms.

In the animal and plant kingdom, ascorbic acid is formed by different routes. In animals, glucose is the primary precursor for the biosynthesis of ascorbic acid, and the last step in the biosynthetic pathway is catalyzed by a microsomal enzyme: L-gulono-γ-lactone oxidase. This enzyme has already been isolated from rat, goat and chicken liver and kidney tissues.

The pathway of ascorbic acid biosynthesis in plants, however, is not yet entirely clear, but there are indications that at least two different biosynthetic pathways exist. Isherwood et al., Biochem. J. 56:1-15 (1954) postulated that the biosynthesis of ascorbic acid starting from D-galactose proceeds via L-galactono-γ-lactone to L-ascorbic acid. Mapson et al., Biochem. J. 56:21-28 (1954) were the first to study this oxidation of L-galactono-γ-lactone to ascorbic acid, a reaction which is catalyzed by L-galactono-γ-lactone dehydrogenase.

The presence of L-galactono-γ-lactone dehydrogenase activity has been described for different plants, including pea, cabbage and potato. Ôba et al., J. Biochem. 117:120-124 (1995) have recently purified the enzyme activity from sweet potato tubers.

Distinct from this biosynthetic pathway, however, an alternative pathway has been proposed which takes as starting point the conversion of D-glucose, and proceeds via L-glucosone and L-sorbosone to ascorbic acid. An NADP-dependent dehydrogenase, which catalyses the conversion of L-sorbosone to ascorbic acid, has been partially purified from bean and spinach leaves (Loewees et al., Plant Physiol. 94:1492-1495 (1990)).

The primary function of ascorbate is as a reducing agent. This is universal. Ascorbic acid is also important as a cofactor for certain enzymatic reactions, including the production of collagen in vertebrates. Since humans are completely dependent on ingested food for the acquisition of ascorbate, it is desirable to increase the vitamin C content of plants and fruit.

Owing to its reducing activity, vitamin C plays a role in the protection of plants and animals against environmental stresses including heat, cold, drought, oxidative stress etcetera. Less stress-sensitive or even stress-resistant plants can therefore play an important part in the economy and agriculture of the world.

It is the object of the present invention to create the possibility of genetically modifying plants such that they contain an increased content of ascorbic acid relative to non-modified plants.

For this purpose the invention provides a polynucleotide in isolated form, which polynucleotide codes for a protein with the activity of the enzyme L-galactono-γ-lactone dehydrogenase, which polynucleotide comprises at least the L-galactono-γ-lactone dehydrogenase activity-determining parts of the coding part of the nucleotide sequence, which is shown in figure 3, or a sequence derived therefrom on the basis of the degeneration of the genetic code. The invention is of course not limited to polynucleotides with exactly the same sequence as that shown in figure 3. It will be apparent to the molecular biologist skilled in the techniques that a certain degree of modification of the

sequence shown in figure 3 is permitted while still falling within the scope of the claim. The polynucleotide is for instance the cDNA shown in figure 3.

Polynucleotides according to the invention can be used in the production of transgenic plant and animal cells, plant tissues or plants with an increased content of the enzyme L-galactono-γ-lactone dehydrogenase relative to non-transgenic plant cells, plant tissues or plants. Such an increased concentration of GLDase will result in plant cells, plant tissues or plants with an increased content of ascorbic acid and with an increased capacity for biosynthesis relative to non-transgenic plant cells, plant tissues or plants.

Plants which can advantageously be used for transformation with the polynucleotide according to the invention are for instance thale cress (Arabidopsis thaliana), tobacco (Nicotiana tabacum), tomato, potato, or corn, without this list being limitative.

Polynucleotides according to the invention can likewise be expressed in eukaryotic cells, such as yeast cells or mammalian cells, in particular fibrosarcoma cells.

The invention further relates to a recombinant L-galactono-γ-lactone dehydrogenase which can be obtained by expression of a polynucleotide according to the invention in a suitable host. The recombinant L-galactono-γ-lactone dehydrogenase can be isolated from transgenic plant tissues or transgenic plants, but also from yeasts or from animal cells.

The invention also relates to a transformation system, comprising a transformation vector or set of vectors, at least one of which includes a nucleotide sequence which codes for the enzyme L-galactono-γ-lactone dehydrogenase. The transformation system preferably comprises Agrobacterium and a binary vector.

Plants or plant tissues with an increased ascorbic acid content can be produced by transforming a plant cell with a gene construct comprising at least the

polynucleotide specified in the invention, optionally linked to targeting sequences for specific organelles, and/or in the presence of suitable transcription and/or translation regulation factors, and regenerating from the plant cell a transgenic plant or plant tissue. The gene construct with the polynucleotide according to the invention can optionally be combined with other genes coding for enzymes which can interfere in the ascorbic acid synthesis, such as L-sorbosone dehydrogenase, UDP-glucuronic acid epimerase, D-galacturonic acid dehydrogenase and ascorbate-regulating enzymes, which may determine the rate of ascorbic acid synthesis.

The enzyme may ultimately be targeted to a particular part of the plant cell, such as the cytoplasm, vacuoles, chloroplasts, mitochondria, lysosomes, endoplasmatic reticulum, Golgi apparatus.

Eukaryotic cells expressing the enzyme GLDase can be obtained by transfection with the polynucleotide according to the invention.

Finally, the invention relates to a new method for purifying the enzyme L-galactono-γ-lactone dehydrogenase. This method comprises of passing a protein extract of cauliflower florets through an ion exchange column: collecting a number fractions eluting from the column and determining the GLDase activity of the fractions; combining fractions with GLDase activity and passing thereof through a Phenyl Sepharose CL 4B column; collecting the column eluate in a number of fractions and determining the GLDase activity of the fractions; combining those fractions with GLDase activity and passing thereof through a gel filtration column; collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions; combining the fractions with GLDase activity and passing through an FPLC Resource Q-column; collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions; combining the fractions with GLDase activity and passing thereof over an FPLC

Poros 20 SP-column; collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions. The enzyme purified by us is lycorine-insensitive, in contrast to the literature which states that L-galactono-γ-lactone dehydrogenase is inhibited by lycorine (De Tullio et al., Boll. Soc. Ital. Biol. Sper. 70:57-62 (1994); Arrigoni et al., Boll. Soc. Ital. Biol. Sper. 72:37-43 (1996)).

Furthermore, the invention provides for a method for increasing the L-ascorbic acid levels in plants, comprising:

a) provision of plants that have been transformed with the sense version of the GLDase gene, and

b) providing the said plants with the precursor L-galactono-γ-lactone in order to induce increased L-ascorbic acid synthesis.

According to another aspect thereof the invention provides transgenic plants having in their genome an antisense version of the GLDase gene resulting in a decreased amount of ascorbic acid as compared to non-transgenic plants for use a model system or biosensor for oxidative stress.

The present invention will be elucidated with reference to the non-limitative examples provided below.

EXAMPLES

EXAMPLE 1

Purification of L-galactono-γ-lactone dehydrogenase

1. Introduction

Using a 5-step purification method which has not previously been described, an acceptable yield of the enzyme L-galactono-γ-lactone dehydrogenase (further designated GLDase) was obtained.

2. Materials and methods

2.1. Materials

Sephacryl SF-200, DEAE Sepharose and Phenyl Sepharose CL-4B were obtained from Pharmacia, Sweden. L-galactono-γ-lactone, D-galactono-γ-lactone, D-gulono-γ-lactone, L-gulono-γ-lactone, L-mannono-γ-lactone, D-galactonic acid, D-glucuronic acid, D-gluconic acid and P-hydroxymercuribenzoic acid were from Sigma Chemical, USA. D-erythronic lactone, D-xylonic lactone and N-ethyl-maleimide were purchased from Aldrich Chemical Company, USA. Restriction enzymes were from Pharmacia, Sweden and (α-^{32}P)dCTP was from Amersham Corp., USA. The cauliflowers (Brassica olecera var. botrytis) were obtained from a field near Ghent and stored at 4°C until use.

2.2. Preparation of an extract

Cauliflower florets (7.5 kg) were cut into small pieces, weighed and homogenized in a pre-cooled blender in ice-cold buffer A (400 mM sucrose, 100 mM sodium phosphate buffer, pH 7.4) (1 l/kg fresh weight). The homogenate was pressed through four layers of Miracloth tissue (Calbiochem-Novabiochem Corp., La Jolla, CA, USA) and centrifuged for 45 minutes at 13,500 x g in a GS3 rotor (Sorvall). The pellet containing the mitochondria (about 250 g material) was kept at -70°C until use.

Before use the pellet was slowly defrosted in a microwave oven and resuspended in 1/10 vol. (750 ml) buffer A. Cold acetone (-20°C) was added slowly while stirring (10 x vol.). The mixture stood for 30 minutes at 4°C. The precipitated protein was collected by filtration through prefilter paper (A15, Millipore, Bedford, USA) and resuspended in 1/10 vol. buffer B (40 mM Tris-HCl, pH 9.0) followed by 5 hours of dialysis against 10 volumes buffer B. The denatured proteins were removed by centrifugation (10,000 x g for 15 minutes). GLDase was purified from the supernatant, further designated as

"protein extract", using the purification procedure described below.

All operations relating to preparation of the extract and enzyme purification were performed at 4°C unless otherwise stated.

2.3. Enzyme purification

The protein extract was placed on a DEAE Sepharose column (5 x 12 cm) equilibrated with buffer B. After washing with 4 volumes of buffer B at a flow rate of 60 ml per hour, the elution was carried out with 0.5 M NaCl in the same buffer. Fractions of 8 ml were collected at a flow rate of 60 ml per hour.

The GLDase activity of the fractions was determined spectrophotometrically by monitoring the L-galactono-γ-lactone dehydrogenase-dependent reduction of cytochrome c at 22°C. A typical reaction mixture contained the enzyme extract, 1.5 mg/ml cytochrome c and 4.2 mM L-galactono-γ-lactone in 0.05 M Tris-HCl buffer (pH 8.4). Reduction of cytochrome c was monitored by determining the absorption increase at 550 nm. Under these conditions the speed of the reaction was linear in respect of time for an initial period of 15 minutes. One unit of enzyme activity was defined as the quantity of enzyme reducing 2 μmol of cytochrome C per minute.

The fractions containing GLDase activity were pooled and ammonium sulphate was added up to a concentration of 1 M. Hereafter the extract was loaded onto a Phenyl Sepharose CL 4B column (2.2 x 15.0 cm) which was equilibrated with buffer C (1 M ammonium sulphate, 25 mM sodium phosphate, pH 7.0). After washing with two volumes of buffer C the elution was carried out with a linear gradient of 0-80% ethylene glycol in 25 mM sodium phosphate, pH 7.0, at a flow rate of 30 ml/hour.

The GLDase activity of the fractions was again determined and GLDase-containing fractions were collected, concentrated to 10 ml by ultrafiltration using a PM 10 membrane (Amicon Corp.) and loaded onto a Sephacryl

SF-200 gel filtration column (2.6 x 94 cm) equilibrated in buffer D (20% ethylene glycol, 40 mM NaCl, 80 mM sodium phosphate, pH 7.4). The enzyme was eluted with the same buffer at a flow rate of 25 ml per hour. Fractions of 5 ml were collected and fractions containing activity were combined. It was possible to keep the gel filtration preparation at 4°C for several weeks without loss of activity.

Two gel filtration preparations were pooled. The preparations were concentrated and the buffer was replaced by buffer E (20% ethylene glycol, 20 mM Tris-HCl, pH 8.0) by means of ultrafiltration. The resulting enzyme solution was loaded onto a 6 ml Resource Q column (Pharmacia) which was equilibrated beforehand with buffer E and coupled to an FPLC system (Pharmacia). The flow rate was 1 ml per minute. Elution was carried out with a gradient of 0 to 450 mM NaCl as follows: 0 to 85 mM in 18 minutes, 85 to 110 mM in 10 minutes, 110 to 130 mM in 14 minutes and 130 to 450 mM in 10 minutes. Fractions of 1 ml were collected. The activity of the main peak, which eluted at 120 mM NaCl, was pooled and brought to pH 6 with 50 mM sodium phosphate.

The pooled fractions were loaded onto a Poros 20 SP column (Pharmacia) coupled to an FPLC and equilibrated in buffer F (20 mM sodium phosphate, pH 6.0, 20% ethylene glycol) at a flow rate of 1 ml/minute. The elution was carried out with a gradient of 0 to 500 mM NaCl in buffer F as follows: 125 to 225 mM in 40 minutes and 225 to 500 mM in 37 minutes. Fractions of 2 ml were collected. Two peaks with activity eluted: peak I at 210 mM and peak II at 225 mM NaCl. Peak II was dialyzed against 10 mM sodium phosphate, pH 7.2.

A Zorbax gel filtration column (9.4 x 250 mm, Zorbax Bioseries GF-250) coupled to an HPLC and equilibrated in 750 mM NaCl, 50 mM sodium phosphate, pH 7.2 was used as final purification step.

Table 1 shows a summary of the purification of GLDase from cauliflower florets. Because the enzymatic

activity was most stable in 20% ethylene glycol, this reagent was included in all buffers except the buffers which were used in the first purifications steps with the DEAE Sepharose and Phenyl Sepharose chromatography. After the DEAE Sepharose step the total GLDase activity increased slightly, probably due to removal of inhibitory compounds which were present in the original crude extract. The FPLC Resource step increased the purification factor from 63 to 900, although the recovery is only 42% in comparison with the activity present in the gel-filtered pool. By the subsequent Poros 20 SP column the activity was separated into two peaks, designated I and II in figure 1. The activity from the latter peak was used for further analysis. Table 1 shows that GLDase was 1693 times more purified from the mitochondrial fraction with a recovery of 1.1%. The purity of the enzyme was tested by means of analytical SDS polyacrylamide gel electrophoresis (SDS PAGE) in slab gels of 10% polyacrylamide as according to Chua, Methods Enzymol. 69:434-446 (1980). Proteins were visualized either by means of Coomassie Brilliant Blue R250 staining (Chua (1980), supra) or silver nitrate staining (Merril et al., Methods Enzymol. 104:441-447 (1984)). Three polypeptide bands were found with molecular masses of about 56 kDa, 30 kDa and 26 kDa (see figure 2).

A partial amino acid sequence was determined as follows. Purified GLDase from the Porous 20 SP purification step was separated by means of SDS-PAGE. The proteins were blotted onto polyvinylidene difluoride (PVDF) membranes (Millipore, Bedford, USA) as described by Bauw et al., Proc. Natl. Acad. Sci. USA 4806-4810 (1987) with 50 mM Tris/50 mM boric acid (pH 8.3) as transfer buffer. The transfer was carried out for at least 8 hours at 35 Volts with a Bio-Rad Transblot apparatus. PVDF membrane-bound polypeptides were visualized by staining with 0.1% Amido black solution. The polypeptide bands were excised and a trypsin digestion was performed in situ, followed by reversed

phase HPLC separation of the generated peptides, as previously described by Bauw et al., Proc. Natl. Acad. Sci. USA 86:7701-7705 (1989). Partial amino acid sequence determination by Edman degradation was carried out on an Applied Biosystems model 473A protein sequencer in accordance with the instructions of the manufacturer.

Table 2 shows the sequences of a number of peptides derived from the GLDase. This shows inter alia that the two low-molecular bands are dissociation products of the 56 kDa band. The NH_2 terminal sequences of the 56 kDa and the 30 kDa polypeptide bands are identical.

EXAMPLE 2

Sensitivity to lycorine

The literature states that lycorine, a pyrrole phenanthridine alkaloid present in different plants of the Amaryllidaceae, inhibits the ascorbic acid synthesis at concentrations from 1 μM. It has recently been demonstrated that the inhibition of lycorine is based on an interaction with the enzyme L-galactono-γ-lactone dehydrogenase (De Tullio et al., (1994), supra; Arrigoni et al., (1996), supra).

Lycorine was isolated from the plant Crinum asiaticum and the identity of the isolated product was verified by NMR, electron spray mass spectrometry HPLC analysis and capillary electrophoresis. Fractions of L-galactono-γ-lactone dehydrogenase activity isolated from the gel filtration column were tested for their activity in two different buffers in the presence of 5 or 50 μM lycorine (see table 3).

All data indicate that the isolated GLDase is insensitive to the inhibitor up to a concentration of 50 μM. Additional tests did not show a decrease in activity even in 100 μM lycorine. A pre-incubation of one hour of the enzyme with lycorine did not influence the enzyme activity.

Table 3
Activity of the enzyme expressed in increase in absorption at 550 nm/second

Concentration lycorine (μM)	0	5	50
75 mM PO4 pH 8	1.882	1.690	1.768
75 mM PO4 pH 7.5	1.385	1.372	1.254
75 mM PO4 pH 7.0	0.980	0.857	0.842
75 mM Tris pH 8.9	5.438	5.199	5.507
75 mM Tris pH 8.2	6.365	6.400	6.127
75 mM Tris pH 7.4	3.627	3.927	3.743

EXAMPLE 3

Isolation of the cDNA

300 mg cauliflower florets were ground to a powder in liquid nitrogen with a pestle and mortar. The powder was suspended in 0.5 ml ice-cold extraction buffer (0.1 M LiCl, 5 mM EDTA, 1% (w/v) SDS and 0.2 M Tris-HCl, pH 7.5) and extracted twice more with phenol/CH_3Cl/isoamyl alcohol (25:24:1). The aqueous phase was adjusted to a final concentration of 3 M LiCl and left on ice for 4 hours. The precipitate was collected by centrifuging for 10 minutes at 20,000 x g and the pellet was washed with 1 ml 3 M LiCl and resuspended in 250 μl H_2O treated with diethyl pyrocarbonate. The LiCl precipitate was repeated and the pellet washed and resuspended in 250 μl H_2O treated with diethyl pyrocarbonate (DEPC). The suspension was centrifuged for 10 minutes at 20,000 x g to remove insoluble material. Sodium acetate was added to an end concentration of 0.3 M followed by addition of 2 volumes ethanol and incubation for 15 minutes at -70°C. The precipitate was collected by centrifuging for 5 minutes at 20,000 x g.

The RNA pellet was washed with 70% ethanol and resuspended in 25 μl H_2O treated with DEPC. The RNA isolated from cauliflower florets (4 μg) was used to synthesize the first strand of cDNA as specified in the in-

struction manual for Superscript™ Preamplification System for First Strand cDNA Synthesis of Gibco BRL.

Degenerated oligonucleotides corresponding with the partial amino acid sequences as shown in example 1 were designed and synthesized on an oligonucleotide synthesizer (Applied Biosystems, Foster City, CA, USA) and used as primers in PCR reactions. The peptide sequences designated 1, 4 and 7 in table 2 were used to design the corresponding coding and complementary oligonucleotides. First-strand cDNA synthesized from cauliflower florets was used as a template. The amplification mixture consisted of 130 ng matrix DNA, PCR buffer (100 mM Tris-HCl, 500 mM KCl, 1.5 mM $MgCl_2$, pH 8.3), 200-300 ng of each primer, 2.5 mM cNTP and 1 unit Taq polymerase in a total volume of 50 μl.

The amplification program consisted of 32 cycles of denaturation for 1 minute at 94°C, annealing for 1 minute at 50°C and primer extension for 2 minutes at 72°C. The reaction products were separated on 1% agarose gels, excised and purified in accordance with the QIAEX handbook of Qiagen GmbH, Germany. The purified products were cloned into a pGEM-T vector (Promega, WI, USA). Of the amplified 250 bp to 400 bp bands which were subcloned into a pGEM-T vector, a 250 bp fragment, which contained a nucleotide sequence corresponding to the amino acid sequence of one of the previously determined internal peptides, was radioactively labelled and used as probe to screen a cDNA library of cauliflower. The cDNA library was constructed in λZAP II (Stratagene, La Jolla, CA, USA) and generously donated by Professor J.S. Hyams (University, London, UK). Portions of the cDNA library were plated using Escherichia coli XL-1 Blue-cells on 23 x 23 cm baking plates (Nunc, Roskilde, Denmark) with NZY agar. About 600,000 plaques from the library were transferred in duplicate to nylon membranes (HYbond N⁺; Amersham Corp., USA). The membranes were treated in accordance with the instructions of the manufacturer for plaque blotting. DNA was fixed to the membranes by

radiation with ultraviolet light (UV Stratalinker, Stratagene, La Jolla, CA, USA). The membranes were subsequently incubated with the 250 bp PCR amplified fragment which was labelled with (α-^{32}P)dCTP with a random primed DNA labelling kit from Boehringer Mannheim, Germany. The membranes were first washed for 4 hours at 65°C in a hybridization buffer (1% (w/v) of bovine serum albumin, 7% (w/v) SDS, 1 mM EDTA and 0.25 M sodium phosphate, pH 7.2) followed by 20 hours of incubation with the ^{32}P-labelled probe in the hybridization buffer at 65°C. The membranes were then rinsed twice for 15 minutes with 2xSSC containing and 1% SDS at room temperature and exposed to X-Omat AR-film (Kodak, CT, USA).

Different positive clones were found. After in vivo excision of the Bluescript plasmid followed by digestion with EcoRI and KpnI the two longest cDNA inserts were found to be approximately 2000 bp long. Subcloning and sequence determination revealed an uninterrupted open reading frame of 1803 nucleotides. The open reading frame contained all the tryptic peptides which had previously been sequenced, the NH_2 terminal amino acid sequence, the first ATG codon (startcodon) (at position 56), and ended with a TAA terminator codon from which it was concluded that the full length cDNA corresponding to the purified protein had been isolated.

Figure 3 shows the derived amino acid sequences of the 1803 bp open reading frame which codes for 600 amino acids. A piece of 55 bp is possibly the 5' non-coding region and a piece of 206 bp shows the 3' non-coding region, including a poly(A)tail. A hexanucleotide AATAAA consensus signal for polyadenylation is found 20 nucleotides before the poly(A)tail. The nucleotides coding for the NH_2 terminal amino acid sequence are found 273 bp from the initiator codon, which indicates that the protein is synthesized as a preprotein (600 amino acids with a calculated molecular mass of 67,829 Da). The resulting mature protein of 509 amino acids has a calculated molecular mass of 57,837 Da and a theoretical

pI-value of 6.85. The number of acidic (Glu and Asp) and basic amino acids (His, Lys and Arg) is respectively 74 and 83. A putative mitochondrial signal peptide is present.

DNA sequence determinations were carried out in accordance with the protocols of US Biochemical Corp. Starting sequences were obtained with the use of T7 and T3 vector primers. Specific primers were used to complete the sequences on both strands of cDNA. The sequence analyses were performed with software from the Genetics Computer Group (Madison, WIC, USA).

EXAMPLE 4

Expression in Yeast

The GLDase cDNA was expressed in Saccharomyces cerevisiae. For this purpose the Bluescript vector containing the complete cDNA was digested with ApaI and KpnI and a 27 bp adaptor containing an NotI restriction site was ligated in the vector linearized with ApaI and KpnI. The resulting construct contains two NotI restriction sites and was cloned in the NotI restriction sites of the pFL61 vector (Minet et al., Plant J. 2:417-422 (1992)). Yeast cells of the strain W303A (Matα, ade 2-1, ura 3-1, his 3-11, 15, trp 1-1, leu 2-3, kan^r) were transformed by means of the method of Dohmen et al., Yeast 7:691-692 (1991) and plated on selective 1.5% agar plates (without uracil) with minimal SD medium (0.2% yeast nitrogen basis (Difco, Detroit, MI, USA), 0.7% ammonium sulphate, 2.7% glucose) supplemented with adenine, tryptophan, leucine at a final concentration of 20 μg /ml, and histidine at a final concentration of 10 μg/ml. Transformed cells were transferred to liquid SD medium (as above but without the agar) and cultured for 3 days at 30°C.

The GLDase cDNA was introduced both in the sense orientation and in the antisense orientation relative to the PGK (phosphoglycerate kinase) promoter and terminator. Non-transformed and transformed yeasts

were grown and extracts were prepared and tested for GLDase activity. Extracts of yeasts which had been transformed with a sense-oriented GLDase cDNA displayed a three- to six-fold increase in specific GLDase activities compared with extracts from non-transformed yeast and yeast which had been transformed with antisense-oriented GLDase cDNA. Wild type yeast has no endogeneous GLDase activity. For determination of protein levels and GLDase activity, cells were harvested by centrifugation (18.000 g, 15 min.), washed and resuspended in 50 mM Tris-HCl (pH 8.0) and disrupted in a French press.

EXAMPLE 5

Transformation of arabidopsis and tobacco

1. Introduction

The GLDase cDNA clone has been used to make sense and antisense GLDase constructs under control of the 35S cauliflower Mosaic Virus (CaMV) promoter. Agrobacterium-mediated transformation has been used to produce transgenic arabidopsis and tobacco plants with the engineered antisense and sense GLDase constructs in order to down-regulate or to up-regulate the GLDase transcript, respectively. Increased GLDase activity was observed in plants transformed with a sense-orientated GLDase cDNA, whereas the specific GLDase activity was low in several antisense plant-lines (see table 4). As a consequence decreased ascorbic acid (AA) levels were measured in antisense transformed plant-lines (see table 5).

2. Materials and methods

2.1. Plasmids and vectors

The GLDase cDNA was inserted in both orientations into the pLBR19 vector (Leple et al. (1992), supra) containing the cauliflower mosaic virus (CaMV) 35S promoter with a double enhancer sequence (CaMV 70). The promoter, enhancer and GLDase cDNA were then cloned into the binary vector pBIN19 (Frisch et al. (1995), supra),

which carries an additional neomycin phosphotransferase (nptII) gene under control of the CaMV 35S promoter.

The sense construct was made as follows: the GLDase cDNA contained in a Bluescript vector was cut with PstI and the resulting partial GLDase cDNA was cloned into the PstI cloning site of the pLBR19 vector in the sense orientation, followed by excision of a SalI-ClaI fragment of this construct. The remaining part of the GLDase cDNA was then inserted as a XhoI-ClaI fragment, resulting in a pLBR19 vector containing the complete GLDase cDNA sequence.

For the antisense construct the following procedure was followed: a fragment of the Bluescript inserted GLDase cDNA was generated by XhoI digestion and inserted into the SalI site of the pLBR19 vector in antisense orientation. Then a SmaI-NsiI fragment was excised from this construct and the remaining part of the GLDase cDNA was inserted as a SmaI-Nsi I fragment resulting in a pLBR19 vector containing the complete GLDase in antisense orientation. The promoter, enhancer, and GLDase cDNA (sense and anti-sense orientation) were finally cloned into the KpnI-XbaI site of the binary vector pBIN19.

The binary plasmids were then mobilized into Agrobacterium, strain C58 Rif (pMP90) as described by Zham et al., Mol. Gen. Genet. 194:188-194 (1984).

DNA electrophoresis, endonuclease digests, ligation reactions and Escherichia coli (strain DH5α) transformations were performed as according to Sambrook et al. (1989), supra.

2.2. Transformation and regeneration

MP90 Agrobacterium tumefaciens (strain C58 Rif) were grown with rifampicin (50 mg/ml), gentamicin (100 mg/l) and kanamycin (200 mg/l) prepared as described by Bechtold et al. (1993), supra and used for plant infection.

2.3. Arabidopsis

Arabidopsis thaliana (columbia cultivar) plants were grown on soil, under standard greenhouse conditions. The plants were transformed by vacuum infiltration as described by Bechtold et al. (1993), supra.

2.4. Tobacco

Transgenic plants were produced from leaf discs of Nicotiana tabacum (SR1) following Agrobacterium-mediated transformation as modified by Thomas et al. (1990), supra. Co-cultivation was for 2-3 days in basal medium (BM) containing 0.5 μM 1-naphtaleneacetic acid and 2.5 μM 6-benzylaminopurine. Leaf discs were then transferred to BM supplemented with the phytohormones mentioned above, and 100 mg/ml kanamycin (Sigma, St. Louis, MO) and 500 mg/l carbenicillin (Sigma). Shoots that formed after 4 weeks were rooted in phytohormone-free BM containing kanamycin.

Plants were transferred to soil, grown under standard greenhouse conditions and self-pollinated. Mature seeds were collected and selected by germination in the presence of kanamycin (125 mg/l).

2.5. Protein extraction

Extracts from plants were prepared by grinding 7 g fresh tissue in liquid nitrogen. Four volumes of buffer containing 100 mM sodium phosphate (pH 7.4) containing 400 mM sucrose were added. The homogenate was squeezed through four layers of Miracloth tissue and centrifuged at 22,000 x g for 30 min. The pellet was resuspended in 5 ml 100 mM sodium phosphate (pH 7.4). Cold acetone (50 ml, -20°C) was slowly added under stirring and the mixture allowed to stand for 30 min. at 4°C. The precipitated protein was collected by centrifugation (10,000 x g for 15 min.). The pellet was dried under vacuum for 30 min. and resuspended in 0.5 ml 40 mM Tris-HCl buffer (pH 8.5). Insoluble proteins were removed by centrifugation (10,000 x g for 15 min.). This

preparation was desalted by gelfiltration on pre-packed NAP-10 (Pharmacia) and used for GLDase activity assays.

2.6. Screening of a genomic library prepared by Arabidopsis thaliana

For screening of a genomic library of Arabidopsis thaliana, the GLDase cDNA was radiolabelled and used as a probe. Five positive clones were isolated. DNA from the largest of these five clones was digested with several restriction enzymes and fractionated on 0.8% (w/v) agarose gel and blotted onto a Hybond-N+ membrane (Amersham, USA) as recommended by Amersham. DNA fragments which hybridized to the GLDase cDNA probe were subcloned into pBluescript KS(+) (Stratagene, USA) and sequenced.

3. Results

3.1. Analysis of plants

Transformed plants were found with the positive (sense) orientation of the GLDase CDNA, and these contained GLDase activity at 2 to 3-fold higher levels, as compared to control plants. In the plants transformed with the GLDase cDNA in a negative (anti-sense) orientation, GLDase activity was approximately 25% of the control plants.

The ascorbic acid levels of 28 antisense GLDase plants were generally lower than the control plants. One plant had 35% AA content compared to the controls and several other plants have values around 50%. The AA levels of the sense GLDase plants were generally higher compared to the controls, with one line attaining 134% of the control.

3.2. Isolation of GLDase gene from Arabidopsis

By screening a genomic Arabidopsis library a 3117 bp DNA clone was isolated. Comparison with the GLDase cDNA sequence isolated from cauliflower indicated that the genomic contained 6 introns. The isolated clone contains 260 bp of the promoter region up-stream to the

first ATG (start) codon. The sequence which corresponds to the last 260 bp from the 3'-end of the GLDase cDNA was not found (Figure 5).

4. Conclusions

The results show the presence of a correctly processed and biologically active GLDase cDNA in the transgenic tobacco plants. It was possible to measure increased GLDase activity levels in plants transformed with GLDase cDNA in the sense orientation. Furthermore, a decreased GLDase activity was measured in plants transformed with the corresponding cDNA in the antisense orientation. In these plants lower ascorbic acid levels were measured.

Leaf disc assays did not conclusively show if transformed plants have changed oxidative stress tolerance.

Table 4

enzyme assays transformed tobacco			
plants	total activity (units/min.)	specific activity (units/min. x mg protein)	%
control	6.0	1.8	100%
sense	5.6	2.1	117%
sense	2.2	3.7	206%
sense	10.5	3.0	167%
antisense	2.6	0.5	28%
antisense	1.3	0.4	22%

Table 5

Anti-oxidant status of transgenic Nicotiana in nmoles/gram fresh weight

plants	L-AA	L-DHA	total
control	1135	132	1267 (100%)
sense	1550	152	1702 (134%)
antisense	520	34	554 (44%)

L-AA = ascorbic acid

L-DHA = oxidized L-AA (dehydro-ascorbic acid)

EXAMPLE 6

Expression in murine fibrosarcoma cells

1. Construction of the eukaryotic expression vector pCAGGS/L-galactono-γ-lactone dehydrogenase

pCAGGS is an expression vector which is used for the efficient expression of genes under the control of the chicken ß-actin/rabbit ß-globin hybrid promoter + CMV-IE enhancer in different mammalian cells (figure 4). The plasmid is a gift from Prof. J. Miyazaki (University or Tokyo, Japan) (Niwa et al., Gene 108:193-200 (1991)).

The L-galactono-γ-lactone dehydrogenase gene was isolated after digestion of the Bluescript SK vector with XbaI and KpnI. The KpnI site was blunted with T4 DNA polymerase and the XbaI/blunt fragment was cloned into the XbaI/BalI sites of the pCAGGS vector. The XbaI site of the pCAGGS is situated at the end of the actual promoter, but the use of this site for cloning a gene has no effect on the expression efficiency.

2. Transfection procedure: stable transfection via DNA calcium phosphate precipitation technique

2.1. Preparation of the cells

28 hours before transfection, L929sA murine fibrosarcoma cells are placed in culture at a concentration of 2.10^6 cells per culture bottle of 75 cm^2. The culture medium used is Dulbecco's modified essential medium (DMEM) enriched with 5% foetal calf serum (FCS), 5% newborn calf serum (NCS), 3 mM glutamine and the antibiotics streptomycin and penicillin. The culture conditions used are 37°C, 5% CO_2.

4 hours before transfection the culture medium is replaced by 10 ml HEPES-buffered minimum essential medium (MEM-HEPES) enriched with 10% FCS, 3 mM glutamine and antibiotics.

2.2. Preparation of DNA precipitate

DNA calcium phosphate precipitate is prepared by adding 30 µg DNA (in 0.25 M $CaCl_2$/0.125 M HEPES

pH 7.05) to the same volume 2x concentrated phosphate/HEPES buffer (0.25 M HEPES, 0.27 M NaCl, 6.7 mM $CaCl_2$, 1.5 mM Na_2HPO_4). The 30 µg DNA is composed from 19 µg carrier DNA (irrelevant plasmid DNA) + 1 µg DNA of the selection plasmid (pSV2 neoplasmid carrying the neomycin resistance gene) + 10 µg pCAGGS/L-galactono-γ-lactone dehydrogenase (plasmid with relevant gene).

2.3. Transfection

The DNA precipitate is placed together with 10 µM chloroquine on the cells, and the mixture incubated for 4 hours in 5% CO_2 at 37°C. The medium with DNA is then removed from the cells and the cells are further held in culture with DMEM.

2.4. Growth and isolation of individual cell colonies

The following day the transfected cells are diluted to a concentration of 250,000 cells per culture bottle of 75 cm^2 and these are further held in culture through selection with the antibiotic G418. After 10-12 days individual colonies can be picked up out of the culture bottle.

The selected colonies are cultured and analyzed for expression of the L-galactono-γ-lactone dehydrogenase. The clones designated with V3, V6, V8 and V14 were found to be positive for the expression of the L-galactono-γ-lactone dehydrogenase. The enzyme activities are shown in table 6.

Table 6

GLDase activity in transfected murine fibrosarcoma cells. All values are expressed in specific activity of the enzyme (units/min./mg protein).

cell line:	VI pod (control)	0
	N2 (control)	0
transfected:	V6	3.7
	V14	2.1
	V3	1.4
	V8	2.0

Table 1. Purification diagram for GLDase

A mitochondrial extract of 15 kg cauliflower florets was used for the preparation.

STEP	VOL. (ml)	PROTEIN (mg)	TOTAL ACTIVITY units	SPECIFIC ACTIVITY units/mg	-FOLD	YIELD %
Acetone prec.	2500	1510	44,900	30.5	1	100
DEAE ion exchanger	83	54.7	46,500	845	28	104
Phenyl Sepharose	38	21.2	30,800	1,467	49	69
Gel filtration	54	10.5	20,900	1,900	63	47
FPLC Resource Q	32	0.3	8,100	2,700	900	18
FPLC Poros 20 SP	4	0.01	508	50,800	1693	1.1

Table 2. Amino acid sequences determined from the GLDase polypeptide

X refers to amino acid sequences not determined by sequence determining runs. Degenerated oligonucleotides were designed on the basis of peptides 1, 4 and 7.

Peptide sequences obtained from 55 kDa polypeptide after tryptic digestion

NH_2-terminal sequences

YAPLXEDL

Internal sequences

LXDQYSAYE (1)

VNQAEAEF (2)

LIALDPLNDVHVG (3)

YTTEEALK (4)

WTGR (5)

GTIELSK (6)

VNQAEAEFWK (7)

IEIPK (8)

Peptide sequences obtained from 31 kDa and 26 kDa subdivisions

NH_2-terminal sequences

APLPDLHTVSN (30 kDa)

XSSKKTPDXRXPDINXL (26 kDa)

CLAIMS

1. Polynucleotide in isolated form, which polynucleotide codes for a protein with the activity of the enzyme L-galactono-γ-lactone dehydrogenase, which polynucleotide comprises at least the L-galactono-γ-lactone dehydrogenase activity-determining parts of the coding part of the nucleotide sequence, which is shown in Fig. 3, or a sequence derived therefrom on the basis of the degeneration of the genetic code.

2. Polynucleotide as claimed in claim 1, which polynucleotide is a cDNA which codes for the enzyme L-galactono-γ-lactone dehydrogenase and at least substantially comprises the coding part of the nucleotide sequence which is shown in Fig. 3.

3. Polynucleotide as claimed in claim 1 or 2 for use in the production of transgenic plant cells, plant tissues or plants with an increased content of the enzyme L-galactono-γ-lactone dehydrogenase relative to non-transgenic plant cells, plant tissues or plants.

4. Polynucleotide as claimed in claim 1 or 2 for use in the production of transgenic plant cells, plant tissues or plants with an increased content of ascorbic acid relative to non-transgenic plant cells, plant tissues or plants.

5. Polynucleotide as claimed in claim 1 or 2 for use in the transformation and/or transfection of eukaryotic cells in order to bring about expression of the polynucleotide therein.

6. Transgenic plant cells which carry in their genome a polynucleotide as claimed in claim 1 or 2 not naturally present therein.

7. Transgenic plant cells as claimed in claim 6 which form part of a transgenic plant tissue and/or a transgenic plant.

8. Transgenic plant tissue consisting at least partially of transgenic plant cells as claimed in claim 6.

9. Transgenic plant tissue as claimed in claim 8 which forms part of a transgenic plant.

10. Transgenic plant which consists at least partially of plant cells as claimed in claim 6.

11. Transgenic plant as claimed in claim 10, obtainable by transforming a plant cell with a polynucleotide as claimed in claim 1 or 2 and by regenerating a plant from the transformed plant cell.

12. Transgenic plant as claimed in claim 10 or 11, **characterized in that** the plant is thale cress (Arabidopsis thaliana), tobacco (Nicotiana tabacum), tomato, potato or corn.

13. Transformed and/or transfected eukaryotic cell, comprising in its genome a polynucleotide as claimed in claim 1 or 2.

14. Recombinant L-galactono-γ-lactone dehydrogenase, obtainable by expression of a polynucleotide as claimed in claim 1 or 2 in a suitable host.

15. Recombinant L-galactono-γ-lactone dehydrogenase as claimed in claim 14 which is isolated from a transgenic plant tissue as claimed in claim 7 or 8, a transgenic plant as claimed in claim 9 or 10, or a eukaryotic cell as claimed in claim 13.

16. Transformation system, comprising a transformation vector or set of vectors, at least one of which includes a nucleotide sequence which codes for the enzyme L-galactono-γ-lactone dehydrogenase.

17. Transformation system as claimed in claim 15, comprising Agrobacterium and a binary vector comprising a polynucleotide as claimed in claim 1 or 2.

18. Use of polynucleotide as claimed in claim 1 or 2 and/or the transformation system as claimed in claim 16 or 17 for producing a transgenic plant or plant tissue with an increased content of L-galactono-γ-lactone dehydrogenase compared with a non-transgenic plant or plant tissue.

19. Use of the polynucleotide as claimed in claim 1 or 2 and/or the transformation system as claimed in claim 16 or 17 for producing a transgenic plant or plant tissue with an increased ascorbic acid content compared to a non-transgenic plant or plant tissue.

20. Use as claimed in claim 17 or 18, wherein the plant is thale cress (Arabidopsis thaliana), tobacco (Nicotiana tabacum), tomato, potato or corn.

21. Use of a polynucleotide as claimed in claim 1 or 2 for transfecting and/or transforming a eukaryotic cell.

22. Method for producing plants or plant tissues with an increased ascorbic acid content, comprising of transformation of a plant cell with a gene construct which comprises at least the polynucleotide as claimed in claim 1 or 2, optionally in the presence of suitable transcription and/or translation regulation factors, and regeneration of a transgenic plant or plant tissue from the plant cell.

23. Method for producing the enzyme L-galactono-γ-lactone dehydrogenase, comprising of transfecting and/or transforming a eukaryotic cell with a gene construct which comprises at least the polynucleotide as claimed in claim 1 or 2, optionally in the presence of suitable transcription and/or translation regulation factors, expressing the enzyme in the transfected and/or transformed cell and optionally isolating the enzyme from the cell and/or its culture medium.

24. Gene construct comprising a polynucleotide as claimed in claim 1 or 2, optionally in the presence of transcription and/or translation regulation factors.

25. Gene construct as claimed in claim 24, further comprising targeting sequence for targeting the encoded enzyme to various parts of the plant cell.

26. Gene construct as claimed in claim 25, wherein the parts of the plant cell are the cytoplasm,

vacuoles, chloroplasts, mitochondria, lysosomes, endoplasmatic reticulum, Golgi apparatus.

27. Method for purifying the enzyme L-galactono-γ-lactone dehydrogenase, comprising of:

a) passing a protein extract of cauliflower florets through an ion-exchange column:

b) collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions;

c) combining the fractions with GLDase activity and passing them through a Phenyl Sepharose CL 4B column;

d) collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions;

e) combining the fractions with GLDase activity and passing them through a gel filtration column;

f) collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions;

g) combining the fractions with GLDase activity and passing them through an FPLC Resource Q-column;

h) collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions;

i) combining the fractions with GLDase activity and passing them through a FPLC Poros 20 SP-column;

j) collecting a number of fractions eluting from the column and determining the GLDase activity of the fractions.

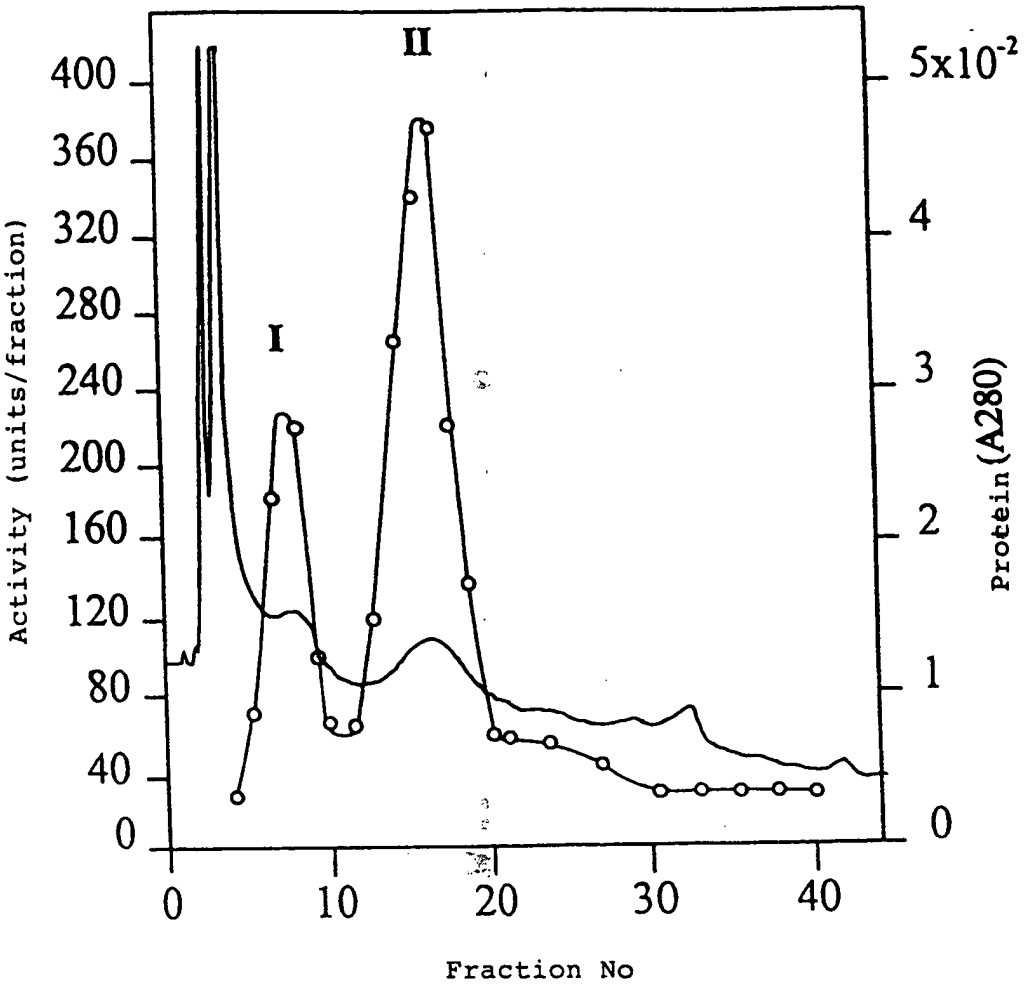
28. Method for increasing the L-ascorbic acid levels in plants, comprising:

a) provision of plants that have been transformed with the sense version of the GLDase gene, and

b) providing the said plants with the precursor L-galactono-γ-lactone in order to induce increased L-ascorbic acid synthesis.

29. Transgenic plants having in their genome an antisense version of the GLDase gene resulting in a decreased amount of ascorbic acid as compared to non-transgenic plants for use a model system or biosensor for oxidative stress.

FIG.1


II
I
400
360
320
280
240
200
160
120
80
40
0
Activity (units/fraction)
5x10-2
4
3
2
1
0
Protein (A280)
0
10
20
30
40
Fraction No

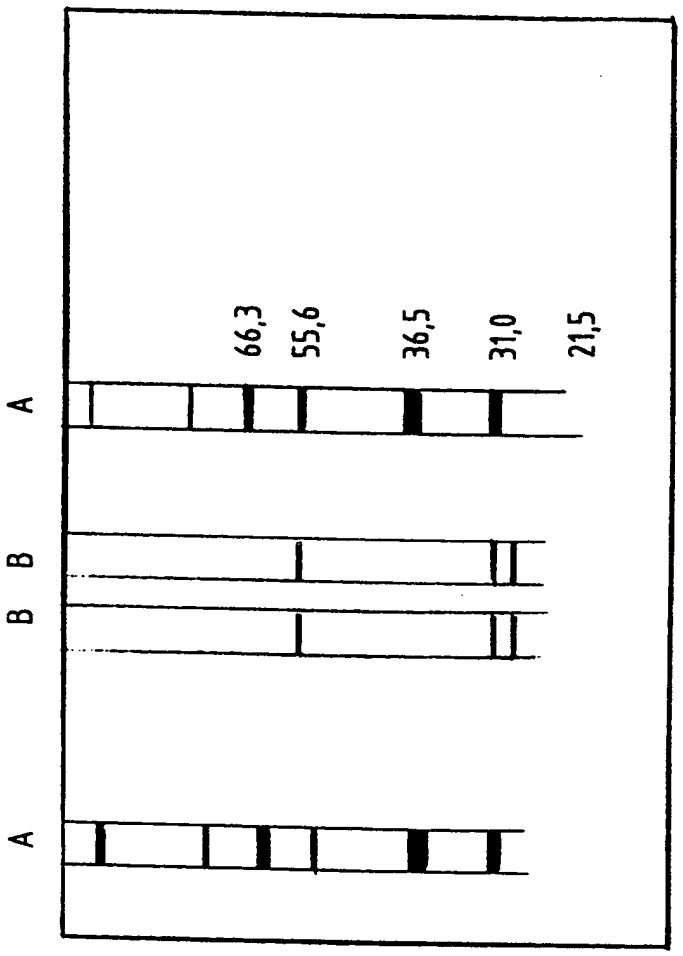


FIG.2

```
1                                                                gctttcgctggctcaggtttcagatcgcctgaactaaaacaaaatg
                                                                                                             M

47   ctccgatcacttctcctccgccgctccaacgcccgttcgcttcgacccccatttccccctctccgcactctatgcacttccggtcagacc
-90    L  R  S  L  L  L  R  R  S  N  A  R  S  L  R  P  P  F  P  P  L  R  T  L  C  T  S  G  Q  T

137  ttgactccagcccctccaccgccgcctcctcctccaccgccgatttcatcctccgcctcagaaaaggagttccgtaaatacgccggatac
-60    L  T  P  A  P  P  P  P  P  P  P  P  P  P  P  I  S  S  S  A  S  E  K  E  F  R  K  Y  A  G  Y

227  gcagcactcgctctcttctccggcgccgcaacttacttctccttcccctttccccgagaacgccaaacacaagaaggctcagatcttccga
-30    A  A  L  A  L  F  S  G  A  A  T  Y  F  S  F  P  F  P  E  N  A  K  H  K  K  A  Q  I  F  R

317  tacgctcctctccccgaagatctccacaccgtctctaactggagtggtactcacgaggtccagaccaggaactttaaccagccggagact
1     Y  A  P  L  P  E  D  L  H  T  V  S  N  W  S  G  T  H  E  V  Q  T  R  N  F  N  Q  P  E  T

407  ctcgccgatctcgaagctctcgtcaaggaagctcatgagaagaagaacaggatccgacccgttggatccggtctttcccccaatgggatc
30    L  A  D  L  E  A  L  V  K  E  A  H  E  K  K  N  R  I  R  P  V  G  S  G  L  S  P  N  G  I

497  ggtttgtctcgctcggggatggtgaatttggcgctcatggacaaggtcctcgaggtggataaagagaagaagagagtccgtgtgcaggct
60    G  L  S  R  S  G  M  V  N  L  A  L  M  D  K  V  L  E  V  D  K  E  K  K  R  V  R  V  Q  A

587  gggattagggttcagcagcttgttgacgccattcaagagtatggtctcactctccagaactttgcttccattagagagcagcagattggt
90    G  I  R  V  Q  Q  L  V  D  A  I  Q  E  Y  G  L  T  L  Q  N  F  A  S  I  R  E  Q  Q  I  G

677  ggcatcattcaggttggggcacatgggacaggtgctagattgcctcctatcgatgagcaagtgattggcatgaagcttgtcactcctgct
120   G  I  I  Q  V  G  A  H  G  T  G  A  R  L  P  P  I  D  E  Q  V  I  G  M  K  L  V  T  P  A

767  aagggaactattgagctttctaaggataatgatccggagctctttcatcttgctcgatgtggccttggtggacttggagttgttgctgag
150   K  G  T  I  E  L  S  K  D  N  D  P  E  L  F  H  L  A  R  C  G  L  G  G  L  G  V  V  A  E

857  gtcaccctccagtgcgttgaaagacaggagcttttggagcacacttacgtctccaccttggaagagatcaagaaaaatcacaaaaagttg
180   V  T  L  Q  C  V  E  R  Q  E  L  L  E  H  T  Y  V  S  T  L  E  E  I  K  K  N  H  K  K  L
```

FIG.3A

```
947  ctctctacaaataagcatgtcaagtacctgtatattccatatactgacacggtcgtggttgttacatgcaaccctgtatcaaaatggagt
210   L  S  T  N  K  H  V  K  Y  L  Y  I  P  Y  T  D  T  V  V  V  V  T  C  N  P  V  S  K  W  S

1037 ggggcacctaaggacaaaccaaagtacactacagaggaggctttaaagcatgtccgtgacctgtatagagagagcattgttaagtatagg
240   G  A  P  K  D  K  P  K  Y  T  T  E  E  A  L  K  H  V  R  D  L  Y  R  E  S  I  V  K  Y  R

1127 gtccaggactctagtaagaagactcctgacagtagggagccagacattaacgagctttcatttacagagttgagagataagctgattgcc
270   V  Q  D  S  S  K  K  T  P  D  S  R  E  P  D  I  N  E  L  S  F  T  E  L  R  D  K  L  I  A

1217 ctagatcctctcaatgacgttcacgttggaaaagtgaatcaagctgaggctgagttttggaaaaaatcagaaggatacagagtagggtgg
300   L  D  P  L  N  D  V  H  V  G  K  V  N  Q  A  E  A  E  F  W  K  K  S  E  G  Y  R  V  G  W

1307 agtgatgaaatcctgggctttgactgtggtggtcaacagtgggtatcagaaacttgttttcctgctggaactctcgctaaacctagcatg
330   S  D  E  I  L  G  F  D  C  G  G  Q  Q  W  V  S  E  T  C  F  P  A  G  T  L  A  K  P  S  M

1397 aaagaccttgagtacatagaacagctgaaagagttgatacaaaaagaagcaataccagcaccttctcccatagagcagcgttggacaggc
360   K  D  L  E  Y  I  E  Q  L  K  E  L  I  Q  K  E  A  I  P  A  P  S  P  I  E  Q  R  W  T  G

1487 cgaagtaagagccctatgagtcctgcattcagcactgcagaggaggacattttctcatgggttggtataatcatgtatctcccgacagca
390   R  S  K  S  P  M  S  P  A  F  S  T  A  E  E  D  I  F  S  W  V  G  I  I  M  Y  L  P  T  A

1577 gaccctcgccagagaaaggacatcacggatgaatttttccactatagacatttgacacaggcaaaattgtgggaccagtattctgcgtat
420   D  P  R  Q  R  K  D  I  T  D  E  F  F  H  Y  R  H  L  T  Q  A  K  L  W  D  Q  Y  S  A  Y

1667 gaacattgggctaaaattgagataccaaaggataaagaggaacttgaagccctacaagaaagactcagaaaacgattcccggtggatgca
450   E  H  W  A  K  I  E  I  P  K  D  K  E  E  L  E  A  L  Q  E  R  L  R  K  R  F  P  V  D  A

1757 tacaacaaagcacgaagggagctggacccaaacagaattctctcaaacaacatggtggaaaagctcttccctgtctccaagactgcttaa
480   Y  N  K  A  R  R  E  L  D  P  N  R  I  L  S  N  N  M  V  E  K  L  F  P  V  S  K  T  A  O
```

FIG.3B

1847 aaacatttcatcaatagttttttgctccttgaagtaccacttttggaatcctataacgttgcatctacaagtgtttgtaagaagagtg

1937 aagccgatatattggtcacaaaaaaagtttacattgagttttactactatttttttttcgcagttcccctgaataaatatacttgttgt

2027 tctattcc

FIG.3C

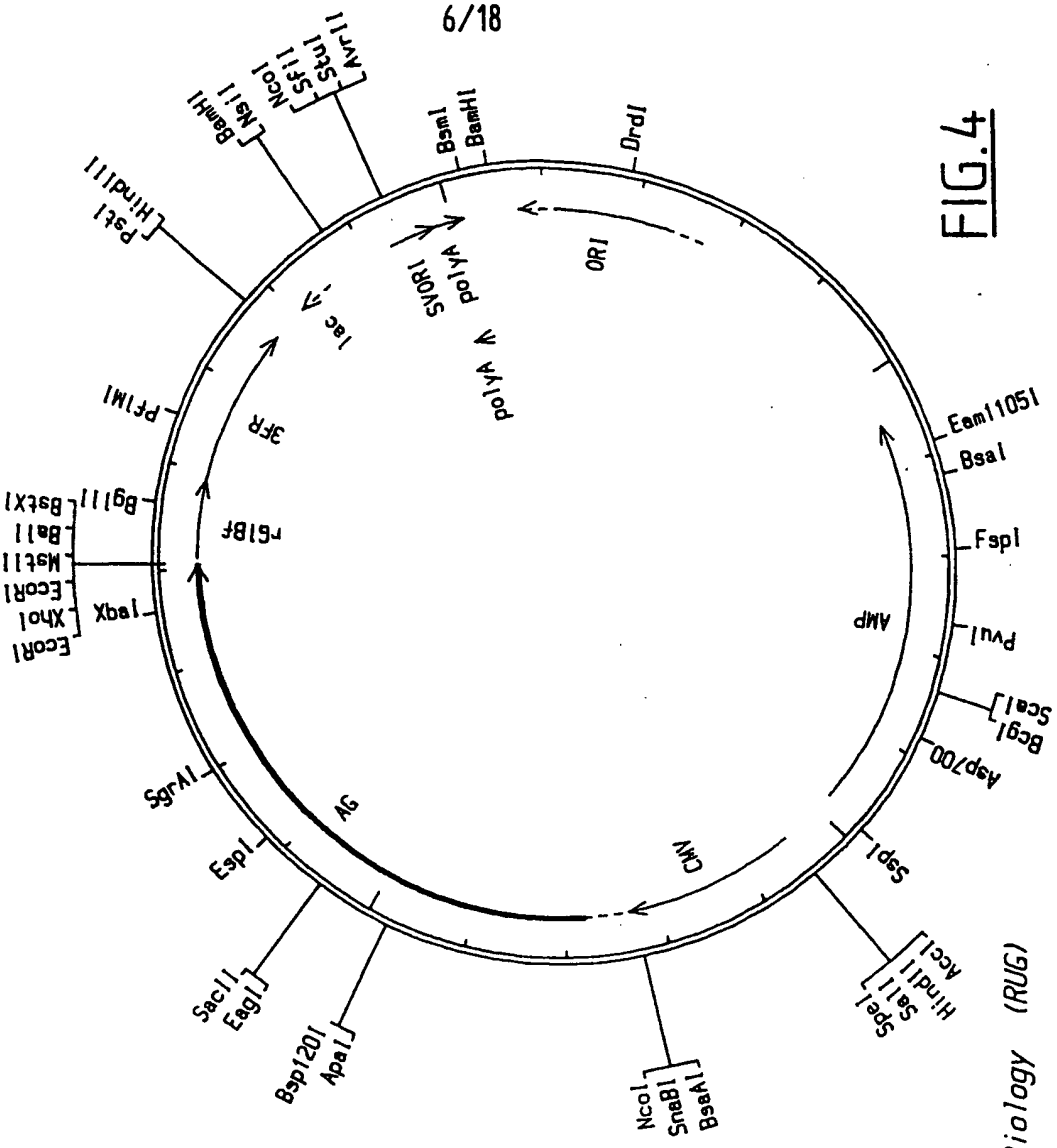

pCAGGS
SN : 2453
max. cutting freq. = 2
35 unique sites
length : 4818 (appr.)
divisions : 1000
subdiv. : 200
(c) Laboratory of Molecular Biology (RUG)
ORI
AMP
CMV
AG
3'R
rG1Bf
lac
SVORI
polyA
BsmI
BamHI
DrdI
EcoRV
PflMI
BglII
BstXI
BalI
MstII
EcoRI
XbaI
XhoI
EcoRI
PstI
HindIII
BamHI
NsiI
NcoI
SfiI
StuI
AvrII
Eam1105I
BsaI
FspI
PvuI
ScaI
BcgI
Asp700
SspI
SpeI
SalI
HindIII
AccI
NcoI
SnaBI
BsaAI
Bsp120I
ApaI
SacI
EagI
EspI
SgrAI

FIG.4

FIG.5A

```
   accgttcgacccgattctcatgcgggacagaaaaccaaaaggcccaaaactacaagtcta
 1 ---------+---------+---------+---------+---------+---------+ 60
   tggcaagctgggctaagagtacgccctgtcttttggttttccgggttttgatgttcagat

   caataaaatttctggttttgtttggtttttgaatgtggacaaactagttaccaatttgtt
61 ---------+---------+---------+---------+---------+---------+ 120
   gttattttaaagaccaaaacaaaccaaaaacttacacctgtttgatcaatggttaaacaa

   cattaacaaattactcggctcaaattatgaaaacagaaataaaatcagggtataatggaa
121 ---------+---------+---------+---------+---------+---------+ 180
   gtaattgtttaatgagccgagtttaatacttttgtctttattttagtcccatattacctt

   actttcttaaatcactaaacccgatcctgtacaagaacatttccctcaggttcagatcgc
181 ---------+---------+---------+---------+---------+---------+ 240
   tgaaagaatttagtgatttgggctaggacatgttcttgtaaagggagtccaagtctagcg

   ctgaagttaaacaaaaaaaaatgctccggtcacttcttctccgacgctccgtcggccattc
241 ---------+---------+---------+---------+---------+---------+ 300
   gacttcaatttgtttttttttacgaggccagtgaagaagaggctgcgaggcagccggtaag

                  M   L   R   S   L   L   L   R   R   S   V   G   H   S   -
```

FIG.5B

```
    tctcggaaccctatctccgtcttcatccaccatccgttcctcattttcgcctcatcgtac
301 ---------+---------+---------+---------+---------+---------+ 360
    agagccttgggatagaggcagaagtaggtggtaggcaaggagtaaaagcggagtagcatg

    L  G  T  L  S  P  S  S  S  T  I  R  S  S  F  S  P  H  R  T  -


    tctctgcaccaccggtcaaacattgacaccaccaccgccgccaccgccacgtcctccacc
361 ---------+---------+---------+---------+---------+---------+ 420
    agagacgtggtggccagtttgtaactgtggtggtggcggcggtggcggtgcaggaggtgg

    L  C  T  T  G  Q  T  L  T  P  P  P  P  P  P  P  R  P  P  P  -


    tccgcctccggccaccgcctcagaagctcaattccgtaaatacgccggatacgcagcact
421 ---------+---------+---------+---------+---------+---------+ 480
    aggcggaggccggtggcggagtcttcgagttaaggcatttatgcggcctatgcgtcgtga

    P  P  P  A  T  A  S  E  A  Q  F  R  K  Y  A  G  Y  A  A  L  -


    cgctatcttctctggagttgctacctatttctcatttccattccctgagaatgctaaaca
481 ---------+---------+---------+---------+---------+---------+ 540
    gcgatagaagagacctcaacgatggataaagagtaaaggtaagggactcttacgatttgt

    A  I  F  S  G  V  A  T  Y  F  S  F  P  F  P  E  N  A  K  H  -
```

FIG.5C

```
    caagaaggctcaaatcttccgttacgctcctttacctgaagatcttcacactgtctctaa
541 ---------+---------+---------+---------+---------+---------+---------+ 600
    gttcttccgagtttagaaggcaatgcgaggaaatggacttctagaagtgtgacagagatt

    K  K  A  Q  I  F  R  Y  A  P  L  P  E  D  L  H  T  V  S  N  -

    ttggagtggtactcatgaggtacagactaggaactttaatcaaccggagaatcttgctga
601 ---------+---------+---------+---------+---------+---------+---------+ 660
    aacctcaccatgagtactccatgtctgatccttgaaattagttggcctcttagaacgact

    W  S  G  T  H  E  V  Q  T  R  N  F  N  Q  P  E  N  L  A  D  -

    tctcgaagctcttgttaaggaatctcatgagaagaagttaaggattcgtcccgttggatc
661 ---------+---------+---------+---------+---------+---------+---------+ 720
    agagcttcgagaacaattccttagagtactcttcttcaattcctaagcagggcaacctag

    L  E  A  L  V  K  E  S  H  E  K  K  L  R  I  R  P  V  G  S  -

    gggtctctcgcctaatgggattggtttgtctcgctctgggatggtgaatctggcgcttat
721 ---------+---------+---------+---------+---------+---------+---------+ 780
    cccagagagcggattaccctaaccaaacagagcgagaccctaccacttagaccgcgaata

    G  L  S  P  N  G  I  G  L  S  R  S  G  M  V  N  L  A  L  M  -
```

FIG.5D

```
    ggataaagttctagaggtggataaagagaagaagagagttacggtgcaggctgggattag
781 ---------+---------+---------+---------+---------+---------+ 840
    cctatttcaagatctccacctatttctcttcttctctcaatgccacgtccgaccctaatc

    D  K  V  L  E  V  D  K  E  K  K  R  V  T  V  Q  A  G  I  R  -


    ggtccagcaattggttgacgccattaaagactatggtcttactcttcagaactttgcctc
841 ---------+---------+---------+---------+---------+---------+ 900
    ccaggtcgttaaccaactgcggtaatttctgataccagaatgagaagtcttgaaacggag

    V  Q  Q  L  V  D  A  I  K  D  Y  G  L  T  L  Q  N  F  A  S  -


    cattagagagcagcagattggtggtattattcaggtttgcatatgtttctctcccttgtg
901 ---------+---------+---------+---------+---------+---------+ 960
    gtaatctctcgtcgtctaaccaccataataagtccaaacgtatacaaagagagggaacac

    I  R  E  Q  Q  I  G  G  I  I  Q  V


    tgaagtctagggttgtgaaactaatggagaatctgaaacaattttagttgttcgtcttta
961 ---------+---------+---------+---------+---------+---------+ 1020
    acttcagatcccaacactttgattacctcttagactttgttaaaatcaacaagcagaaat
```

FIG.5E

```
     tcttgtgctttgaggttttagagtctatattttgtttacgttcaggttggggcacatg
1021 ---------+---------+---------+---------+---------+---------+---------+ 1080
     agaacacgaaactccaaaaatctcagatataaaaacaaatgcaagtccaaccccgtgtac

                                     F   V   Y   V   Q   V   G   A   H   G -

     ggacaggtgctagattgcctcctattgatgagcaggtgatcagtatgaagctggttactc
1081 ---------+---------+---------+---------+---------+---------+---------+ 1140
     cctgtccacgatctaacggaggataactactcgtccactagtcatacttcgaccaatgag

       T   G   A   R   L   P   P   I   D   E   Q   V   I   S   M   K   L   V   T   P -

     ctgCgaagggAAcaattgaactttcaagagagaaagatccggagctctttcatctagctc
1141 ---------+---------+---------+---------+---------+---------+---------+ 1200
     gacGcttcccTTgttaacttgaaagttctctctttctaggcctcgagaaagtagatcgag

       A   K   G   T   I   E   L   S   R   E   K   D   P   E   L   F   H   L   A   R -

     gatgtGGCCttggtGGacttGGagttgttgctgaggtcaccctccaatgcgttgcaagac
1201 ---------+---------+---------+---------+---------+---------+---------+ 1260
     ctacaCCGGaaccaCCtgaaCCtcaacaacgactccagtgggaggttacgcaacgttctg

       C   G   L   G   G   L   G   V   V   A   E   V   T   L   Q   C   V   A   R   H -
```

FIG.5F

```
     atgaacttgtggaacacacatacgtctcaaacttgcaagaaatcaagaaaaatcacaagt
1261 ---------+---------+---------+---------+---------+---------+ 1320
     tacttgaacaccttgtgtgtatgcagagtttgaacgttctttagttctttttagtgttca

       E   L   V   E   H   T   Y   V   S   N   L   Q   E   I   K   K   N   H   K

     taagtatcgctaactttcgctatattagtctccatattatggctccagcttagaaaatca
1321 ---------+---------+---------+---------+---------+---------+ 1380
     attcatagcgattgaaagcgatataatcagaggtataataccgaggtcgaatctttagt

     tgctcagtatacgacttttcttctggtcagattatcagagaagtatcagattgatgcaag
1381 ---------+---------+---------+---------+---------+---------+ 1440
     acgagtcatatgctgaaaagaagaccagtctaatagtctcttcatagtctaactacgttc

     agcttaaagtttttttcacttttagtactgcccatatcatttggcatagtgcattctag
1441 ---------+---------+---------+---------+---------+---------+ 1500
     tcgaatttcaaaaaaagtgaaaaatcatgacgggtatagtaaaccgtatcacgtaagatc

     catttgggaaatcactccctctactttgaagcaaattgatcccataaattggctcagggt
1501 ---------+---------+---------+---------+---------+---------+ 1560
     gtaaacccttttagtgagggagatgaaacttcgtttaactagggtatttaaccgagtccca
```

FIG.5G

```
      ggaacgtttcctaacttttgttttgtttctGGctgttcagaaaattgctctctgcAaaca
1561  ---------+---------+---------+---------+---------+---------+---------+ 1620
      ccttgcaaaggattgaaaacaaaacaaagaCCgacaagtcttttaacgagagacgTttgt

                                                  K   L   L   S   A   N   K -

      agcatgttaagtacctatatattccTtataCcgacacagtcgtggttgtaacatgcaatc
1621  ---------+---------+---------+---------+---------+---------+---------+ 1680
      tcgtacaattcatggatatataaggAatatGgctgtgtcagcaccaacattgtacgttag

        H   V   K   Y   L   Y   I   P   Y   T   D   T   V   V   V   V   T   C   N   P -

      ctgtatcaaaatggagtgggccACCtaaggACAaaccaaagtacactacagatgaggctg
1681  ---------+---------+---------+---------+---------+---------+---------+ 1740
      gacatagttttacctcacccggTGGattccTGTttggtttcatgtgatgtctactccgac

        V   S   K   W   S   G   P   P   K   D   K   P   K   Y   T   T   D   E   A   V -

      tAcagcatgtccgtgatctctacagagagagcattgtgaagtataggtatcgttatgctt
1741  ---------+---------+---------+---------+---------+---------+---------+ 1800
      aTgtcgtacaggcactagagatgtctctctcgtaacacttcatatccatagcaatacgaa

        Q   H   V   R   D   L   Y   R   E   S   I   V   K   Y   R
```

FIG.5H

```
    aagtcttatgtgtaacttgatttctctaatgtggaggactgaatgaaatgcaaaataatt
1801 ---------+---------+---------+---------+---------+---------+ 1860
    ttcagaatacacattgaactaaagagattacacctcctgacttactttacgttttattaa

    ttttactatgatgtatagggtccaggactctggtaagaagtctcctgacagcagtgagcc
1861 ---------+---------+---------+---------+---------+---------+ 1920
    aaaatgatactacatatcccaggtcctgagaccattcttcagaggactgtcgtcactcgg

                    R  V  Q  D  S  G  K  K  S  P  D  S  S  E  P  -

    agacatacaggagctttcatttacagagttgagagacaaactacttgcccttgatcctct
1921 ---------+---------+---------+---------+---------+---------+ 1980
    tctgtatgtcctcgaaagtaaatgtctcaactctctgtttgatgaacgggaactaggaga

     D  I  Q  E  L  S  F  T  E  L  R  D  K  L  L  A  L  D  P  L  -

    caatgacgttcacgttgcaaagtaaatcaagctgaggcagagttttggaaaaaatcagaa
1981 ---------+---------+---------+---------+---------+---------+ 2040
    gttactgcaagtgcaacgtttcatttagttcgactccgtctcaaaacctttttttagtctt

     N  D  V  H  V  g  K  V  N  Q  A  E  A  E  F  W  K  K  S  E  -
```

FIG.5I

```
     ggatatagagtagggtggagtgatgaaattctgggctttgactgtggtggtcagcagtgg
2041 ---------+---------+---------+---------+---------+---------+---------+ 2100
     cctatatctcatcccacctcactactttaagacccgaaactgacaccaccagtcgtcacc

      G   Y   R   V   G   W   S   D   E   I   L   G   F   D   C   G   G   Q   Q   W   -

     gtgtcagaatcttgttttcctgctggaaccctcgccaaccctagcatgaaagaccttgaa
2101 ---------+---------+---------+---------+---------+---------+---------+ 2160
     cacagtcttagaacaaaaggacgaccttgggagcggttgggatcgtactttctggaactt

      V   S   E   S   C   F   P   A   G   T   L   A   N   P   S   M   K   D   L   E   -

     tacatagaagagctgaaaaaactaatagaaaaggaagcaataccagcacctgctccaata
2161 ---------+---------+---------+---------+---------+---------+---------+ 2220
     atgtatcttctcgacttttttgattatcttttccttcgttatggtcgtggacgaggttat

      Y   I   E   E   L   K   K   L   I   E   K   E   A   I   P   A   P   A   P   I   -

     gagcagcgatggacagctcgaagtaagagccccattagtcctgcATTcagcacttcagag
2221 ---------+---------+---------+---------+---------+---------+---------+ 2280
     ctcgtcgctacctgtcgagcttcattctcggggtaatcaggacgTAAgtcgtgaagtctc

      E   Q   R   W   T   A   R   S   K   S   P   I   S   P   A   F   S   T   S   E   -
```

FIG.5J


gatgatattttctcatgggtaactcttgttttatgtcgtttatccttccatttacttctc
2281 ---------+---------+---------+---------+---------+---------+---------+ 2340
ctactataaaagagtacccattgagaacaaaatacagcaaataggaaggtaaatgaagag
D D I F S W V
tttgactttcatgaaagtatgaagagatattggtgtcaatctataggaagcttgttttgt
2341 ---------+---------+---------+---------+---------+---------+---------+ 2400
aaactgaaagtactttcatacttctctataaccacagttagatatccttcgaacaaaaca
ggctctgcctttgtggtggaggaaaacatgtgatatattgatgttaaaatgttcatagac
2401 ---------+---------+---------+---------+---------+---------+---------+ 2460
ccgagacggaaacaccacctcctttgtacactatataactacaattttacaagtatctg
aaagaagaaaccgtaaaaatgatgttacatactgtactcttaggtgctggattgttgttt
2461 ---------+---------+---------+---------+---------+---------+---------+ 2520
tttcttctttggcatttttactacaatgtatgacatgagaatccacgacctaacaacaaa
cacttggtagattttttgttgttggccaaccttgttccaacaccgactgtttgccttttt
2521 ---------+---------+---------+---------+---------+---------+---------+ 2580
gtgaaccatctaaaaacaacaaccggttggaacaaggttgtggctgacaaacggaaaaaa

FIG.5K

```
     ctctttcaaatgctagtcatctacagttataatatgctacattacatttgtctcaggttg
2581 ---------+---------+---------+---------+---------+---------+---------+ 2640
     gagaaagtttacgatcagtagatgtcaatattatacgatgtaatgtaaacagagtccaac

                                                          V   G   -

     gtataatcatgtacctcccgacagcagaccctcgccagagaaaggacatcacagatgaat
2641 ---------+---------+---------+---------+---------+---------+---------+ 2700
     catattagtacatggagggctgtcgtctgggagcggtctctttcctgtagtgtctactta

       I   I   M   Y   L   P   T   A   D   P   R   Q   R   K   D   I   T   D   E   F   -

     ttttccactatagacatttgacacagaaacaattgTGGGatcaaTTTTctgcgtatgaac
2701 ---------+---------+---------+---------+---------+---------+---------+ 2760
     aaaaggtgatatctgtaaactgtgtctttgttaacACCCtagttAAAAgacgcatacttg

       F   H   Y   R   H   L   T   Q   K   Q   L   W   D   Q   F   S   A   Y   E   H   -

     attgggctaaaatTgaggtaatcgtagattttctaatctaaatatgagatTcttgTatct
2761 ---------+---------+---------+---------+---------+---------+---------+ 2820
     taacccgattttaActccattagcatctaaaagattagatttatactctaAgaacAtaga

       W   A   K   I   E
```

FIG.5L

```
     taacacacagataccatcattctcacttaactatgTccttctgattcactcacaaaaagt
2821 ---------+---------+---------+---------+---------+---------+---------+ 2880
     attgtgtgtctatggtagtaagagtgaattgatacAggaagactaagtgagtgtttttca

     ctctgTatcttaattacatttttctgcttgaactacaactgtcctcattgtgagaagta
2881 ---------+---------+---------+---------+---------+---------+---------+ 2940
     gagacAtagaattaatgtaaaaaagacgaacttgatgttgacaggagtaacactcttcat

     agcaaagggaatgagaatctgttgaggtaactatttagagtgtagacaatttctaatgtt
2941 ---------+---------+---------+---------+---------+---------+---------+ 3000
     tcgtttcccttactcttagacaactccattgataaatctcacatctgttaaagattacaa

     ttctgtttgatatttatataatcagataccaaaagacaaagaagaacttgaagccttaca
3001 ---------+---------+---------+---------+---------+---------+---------+ 3060
     aagacaaactataaatatattagtctatggttttctgtttcttcttgaacttcggaatgt

                                   I   P   K   D   K   E   E   L   E   A   L   Q   -

     ggcaagaataagaaaacgtttcccagtggatgcatacaacaaattcctaggatcctcgag
3061 ---------+---------+---------+---------+---------+---------+---------+ 3120
     ccgttcttattcttttgcaaagggtcacctacgtatgttgtttaaggatcctaggagctc

      A   R   I   R   K   R   F   P   V   D   A   Y   N   K
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/53, 15/82, 9/02, A01H 5/00

A3

(11) International Publication Number: **WO 98/50558**

(43) International Publication Date: 12 November 1998 (12.11.98)

(21) International Application Number: PCT/EP98/02830

(22) International Filing Date: 7 May 1998 (07.05.98)

(30) Priority Data:
1006000 7 May 1997 (07.05.97) NL

(71) Applicant *(for all designated States except US)*: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE [BE/BE]; Rijvisschestraat 118, bus 1, B–9052 Zwijnaarde–Poperinge (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BAUW, Guy, Jerome, Corneel [BE/BE]; Blokweg 23, B–8972 Proven–Poperinge (BE). DAVEY, Mark, William [GB/BE]; Patijntjestraat 68, B–9000 Gent (BE). OSTERGAARD, Jens [DK/BE]; Ekkergemstraat 165, B–9000 Gent (BE). VAN MONTAGU, Marc, Charles, Ernest [BE/BE]; De Stassartstraat 120, B–1050 Brussel (BE).

(74) Agent: VAN SOMEREN, Petronella, Francisca, Hendrika, Maria; Arnold & Siedsma, Sweelinckplein 1, NL–2517 GK The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
4 February 1999 (04.02.99)

(54) Title: PRODUCTION OF ASCORBIC ACID IN PLANTS

(57) Abstract

The present invention relates to a polynucleotide in isolated form, which polynucleotide codes for a protein with the activity of the enzyme L–galactono–γ–lactone dehydrogenase, which polynucleotide comprises at least the L–galactono–γ–lactone dehydrogenase activity–determining parts of the coding part of the nucleotide sequence, which is shown in Fig. 3, or a sequence derived therefrom on the basis of the degeneration of the genetic code. The invention further relates to the use of the polynucleotide in the production of transgenic plants, plant cells or other eukaryotic cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/EP 98/02830

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/53 C12N15/82 C12N9/02 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category '	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 97 04100 A (FRAUNHOFER GES FORSCHUNG ;WISSLER JOSEF (DE)) 6 February 1997 see the whole document ---	1-5, 14-16,27 23,24
Y	YAGI, ET AL.: "Expression in monkey cells of the missing enzyme in L-ascorbic acid biosynthesis, L-glucono-gamma-lactone oxidase" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 177, no. 2, 1991, pages 659-663, XP002054716 see the whole document --- -/--	23,24

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

' Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 October 1998	26/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Maddox, A

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. Ional Application No
PCT/EP 98/02830

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MINOBE, Y., ET AL.: "Rice cDNA, partial sequence (R0642_1A)" EMBL DATABASE, REL.37 29-NOV-1993, ACCESSION NO. D23947, XP002054717 see the whole document	1
P,X	OSTERGAARD, J., ET AL.: "Isolation of a cDNA coding for L-galactono-gamma-lactone dehydrogenase, an enzyme involved in the biosynthesis of ascorbic acid in plants" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, 28 November 1997, pages 30009-30016, XP002054719 see the whole document	1-5,14, 15,27
A	BAI Y ET AL: "L-GALACTONO-1,4-LACTONE: FERRICYTICHROME-C OXIDOREDUCTASE E.C.1.3.2.3 FROM PEAS AND ITS USE IN A BIOPROCESS FOR L-ASCORBIC ACID SYNTHESIS" BIOLOGICAL CHEMISTRY HOPPE-SEYLER, vol. 373, no. 9, 25 September 1992, page 857/858 XP002025678 see the whole document	1-29
A	OBA, K., ET AL.: "Purification and properties of L-galactono-gamma-lactone dehydrogenase, a key entyme for ascorbic acid biosynthesis, from sweet potato roots." J. BIOCHEM., vol. 117, 1995, pages 120-124, XP002054718 see the whole document	1-29
A	CHEMICAL ABSTRACTS, vol. 125, no. 15, 1996 Columbus, Ohio, US; abstract no. 190777, ISHII, G., ET AL.: "Biochemical approach to higher ascorbic acid levels in potato breeding" XP002080811 see abstract & SPEC. PUBL. - R. SOC. CHEM., vol. 179, 1996, pages 63-66,	1-29
A	KULBE K D ET AL: "ENZYMATIC SYNTHESIS OF L-ASCORBIC ACID VIA D-URONIC ACIDS: MEMBRANE-REACTOR INTEGRATED RECOVERY OF D-GALACTURONIC ACID FROM PECTIN HYDROLYSATES" ANNALS OF THE NEW YORK ACADEMY OF SCIENCES, vol. 506, 1987, pages 543-551, XP000646778 see page 543 - page 544	1-29

-/--

1

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/EP 98/02830

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WO 85 01745 A (KRAFT INC) 25 April 1985 see the whole document	1-29
A	US 4 259 443 A (DANEHY JAMES P) 31 March 1981 see column 9, last paragraph	1
A	WO 96 31115 A (UNIV CALIFORNIA) 10 October 1996 see page 8, line 10 - line 22 see page 23, line 3 - line 9	6-12, 17-22
A	WO 96 17944 A (CIBA GEIGY AG ;PATTON DAVID ANDREW (US)) 13 June 1996 see the whole document	6-12, 17-22
A	CHEMICAL ABSTRACTS, vol. 124, no. 9, 1995 Columbus, Ohio, US; abstract no. 115786, FUKADA,M., ET AL.: "Effect of storage time of potatoes after harvest on increase in the ascorbic acid content by wounding" XP002054720 see abstract & NIPPON SHOKUHIN KAGAKU KOGAKU KAISHI, vol. 42, no. 12, 1995, pages 1031-1034,	6-12, 17-22
A	WO 93 19193 A (BIO TECH RESOURCES) 30 September 1993 see the whole document	6-13
A	DATABASE WPI Section Ch, Week 8920 Derwent Publications Ltd., London, GB; Class B04, AN 89-148142 XP002054721 & JP 01 091785 A (VITAMIN KENKYUSHO K) , 11 April 1989 see abstract	16

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte .onal Application No
PCT/EP 98/02830

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9704100 A	06-02-1997	AU 6731796 A	18-02-1997
		DE 19628882 A	27-02-1997
WO 8501745 A	25-04-1985	US 4595659 A	17-06-1986
		DK 280285 A	20-06-1985
		EP 0146239 A	26-06-1985
		JP 61500201 T	06-02-1986
		US 4916068 A	10-04-1990
US 4259443 A	31-03-1981	NONE	
WO 9631115 A	10-10-1996	US 5750872 A	12-05-1998
WO 9617944 A	13-06-1996	AU 4300196 A	26-06-1996
		CA 2205561 A	13-06-1996
		EP 0796337 A	24-09-1997
		FI 972336 A	06-06-1997
WO 9319193 A	30-09-1993	AU 3811793 A	21-10-1993
		CA 2091919 A	19-09-1993
		CN 1080955 A	19-01-1994
		JP 6046834 A	22-02-1994
		US 5521090 A	28-05-1996
		US 5792631 A	11-08-1998